In F. Lauze, Y. Dong, A. B. Dahl (Eds.):
Scale Space and Variational Methods in Computer Vision.
Lecture Notes in Computer Science, Vol. 10302, 590-601, Springer, Cham, 2017.
The final publication is available at link.springer.com

Evaluating Data Terms for Variational Multi-frame Super-resolution

Kireeti Bodduna and Joachim Weickert

Mathematical Image Analysis Group,
Faculty of Mathematics and Computer Science,
Saarland University, 66041, Saarbrücken, Germany.
{bodduna,weickert}@mia.uni-saarland.de

Abstract. We present the first systematic evaluation of the data terms for multi-frame super-resolution within a variational model. The various data terms are derived by permuting the order of the blur-, downsample-, and warp-operators in the image acquisition model. This yields six different basic models. Our experiments using synthetic images with known ground truth show that two models are preferable: the widely-used warp-blur-downsample model that is physically plausible if atmospheric blur is negligible, and the hardly considered blur-warp-downsample model. We show that the quality of motion estimation plays the decisive role on which of these two models works best: While the classic warp-blur-downsample model requires optimal motion estimation, the rarely-used blur-warp-downsample model should be prefered in practically relevant scenarios when motion estimation is suboptimal. This confirms a widely ignored result by Wang and Qi (2004). Last but not least, we propose a new modification of the blur-warp-downsample model that offers a very significant speed-up without substantial loss in the reconstruction quality.

Keywords: super-resolution, variational methods, inverse problems

1 Introduction

Generating high resolution (HR) images is one of the main objectives of photography. These images show more details of the scene which is crucial for many applications such as surveillance, medical or satellite imaging. Instead of opting for expensive high precision optics, the other way is to enhance the resolution after capturing. While there exist methods to enhance the resolution of a single image [9, 12, 23], which is referred to as single-frame super-resolution, a more

common technique is to acquire several low resolution (LR) images of the same scene and combine the information from them into a single HR image. This process is known as super-resolution (SR) reconstruction.

Super-resolution is an inverse problem where we are trying to recover the unknown HR image. Given N LR images $\left\{\boldsymbol{f}_L^k\right\}_{k=1}^N$ of resolution $N_L = L_1 \times L_2$, we want to find a HR image $\boldsymbol{f}_H$ of resolution $N_H = H_1 \times H_2$ where $N_H = N_L \times z$ (z is the zoom factor). This HR image $\boldsymbol{f}_H$ should minimize the following energy function :

$$E(\boldsymbol{f}_H) = \sum_{k=1}^{N} |\boldsymbol{P}_k(\boldsymbol{f}_H) - \boldsymbol{f}_L^k|^2 \tag{1}$$

where $\boldsymbol{P}_k(\boldsymbol{f}_H)$ is the projection of the HR image onto the LR scale and $|.|$ denotes the Euclidean norm. The projection operator $\boldsymbol{P}_k(\boldsymbol{f}_H)$ is modeled by a sequence of linear transformations: Blurring, motion warping and downsampling are the three operators that describe the relation between the HR scene and the LR realizations of the scene. The order of these operators is decided keeping in mind the image acquisition procedure. The procedure of acquiring digital images is modeled as follows [8]:

$$\boldsymbol{f}_L^k = \boldsymbol{D}\boldsymbol{B}_{cam}\boldsymbol{W}_k\boldsymbol{B}_{atmp}\boldsymbol{f}_H + \boldsymbol{e}_k \tag{2}$$

The HR scene $\boldsymbol{f}_H$ undergoes an atmospheric blurring matrix $\boldsymbol{B}_{atmp}$ (size: $N_H \times N_H$) first. The objects in the scene are assumed to be moving with the camera as the frame of reference. This motion is modeled by a warping operator $\boldsymbol{W}_k$ (size: $N_H \times N_H$). The moving objects then undergo a blurring due to the point spread function (PSF) of the camera is modeled by $\boldsymbol{B}_{cam}$ (size: $N_H \times N_H$). Finally the HR scene is downsampled by the camera detector system which is modeled with the help of the operator $\boldsymbol{D}$ (size: $N_L \times N_H$). An additive noise vector $\boldsymbol{e}_k$ of size N_L is assumed. This entire process leads to a single LR realization $\boldsymbol{f}_L^k$ of size N_L. It should be mentioned that the blurring and the downsampling operators do not have an index k as the atmospheric conditions and the camera conditions are assumed to be same for all the LR realizations.

The seminal work on multi-frame super-resolution goes back to Tsai et al. [20]. The standard observational model for SR reconstruction that is widely followed [1, 2, 5, 7, 10, 11, 13–16, 16, 18, 19, 24] was first proposed and represented in a matrix-vector formulation by Elad et al. [6]:

$$\boldsymbol{f}_L^k = \boldsymbol{D}\boldsymbol{B}\boldsymbol{W}_k\boldsymbol{f}_H + \boldsymbol{e}_k \tag{3}$$

One can see that the atmospheric blurring operator from Equation 1 has been dropped. Dealing with an extra blurring matrix might be a bit cumbersome both with respect to the number of parameters in the model and the computational burden. Also, on the other hand it is not a bad assumption to assume that the camera blur is the dominant one amongst the two type of blurs, especially in cases where images are obtained from microscopes.

Our contribution. Evaluating the SR observational models that either deviate or do not deviate from the image acquisition process in terms of ordering of the operators, is still an open research field. Very few works [18, 21, 25] have concentrated their research on this particular problem. Variational models are known for the many degrees of freedom they provide while modeling various problems. In this work, we make use of this liberty provided by variational models to systematically evaluate the different SR data terms. These different SR data terms are derived by modifying the image acquisition model but are mathematically still plausible. We only focus on the data term as the discussions for the smoothness terms are similar to that of other applications like image denoising and optic flow. Finally, we also propose a new modified observational model which helps to save a lot of computational time.

Paper structure. The outline of this paper is as follows: We propose various SR observational models and introduce the optic flow method used in calculating the warping matrix in Section 2. The experiments and discussions on the results from the experiments are presented in Section 3. We consolidate the conclusions from the performed experiments and discussions in Section 4.

2 Modeling and Theory

2.1 Optic Flow

The warping matrix in Equation 2 represents the displacements that the objects in the HR scene have undergone before being captured as a LR image by the camera. We make use of a simplified version of the popular optic flow method by Brox et al. [4] to estimate this matrix (we omit gradient constancy and just consider grey value constancy). This method is designed specifically for handling large displacements by using a theory of multi-scale warping. Also, the method does not assume a particular type of motion and hence it is a very good fit for estimating the warping matrix. We consider one of the LR images to be the reference image. The warping matrix is calculated for every reference image - LR image pair. In the following we briefly sketch the main ideas behind this optic flow method.

Let $\boldsymbol{x} := (x, y, t)^T$ denote the position vector and $\boldsymbol{w} = (u, v, 1)^T$ the unknown displacement vector field. Penalizing the deviations from the grey value constancy and enabling interaction between pixels can be modeled by the following continuous energy functional :

$$E(u,v) = \int_{\Omega} \left(\Psi(|f(\boldsymbol{x}+\boldsymbol{w}) - f(\boldsymbol{x})|^2 + \alpha_{OF}(\Psi(|\boldsymbol{\nabla} u|^2 + |\boldsymbol{\nabla} v|^2) \right) d\boldsymbol{x} \quad (4)$$

where Ω is the image domain, f denotes the image sequence and $\boldsymbol{\nabla}$ represents a spatio-temporal gradient. To tackle the effect of outliers on a quadratic energy, an increasing convex function $\Psi(s^2)$ is applied for a robust convex energy functional such as $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ with a small positive constant ϵ required

Table 1: Proposed SR observational models along with the mean squared error (MSE) values of the obtained reconstructed SR image for TS1 using flows F1 and F2. The optimized parametric values are also shown.

#	Observational Model	MSE (F1)	MSE (F2)	σ (F1)	σ (F2)	α (F1)	α (F2)
1	$\boldsymbol{DBW}_k\boldsymbol{f}_H = \boldsymbol{f}_L^k$	**103.26**	170.11	0.57	0.59	0.002	0.006
2	$\boldsymbol{DW}_k\boldsymbol{B}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	121.76	171.27	0.65	0.61	0.004	0.005
3	$\boldsymbol{BDW}_k\boldsymbol{f}_H = \boldsymbol{f}_L^k$	120.63	193.23	0.63	0.65	0.009	0.06
4	$\boldsymbol{W}_k\boldsymbol{DB}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	116.46	164.15	0.57	0.51	0.0008	0.0004
5	$\boldsymbol{BW}_k\boldsymbol{D}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	120.12	173.93	0.58	0.55	0.0004	0.0007
6	$\boldsymbol{W}_k\boldsymbol{BD}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	121.91	173.51	0.61	0.56	0.0008	0.0005

for retaining the convex property of the energy functional after the application of Ψ. Moreover, α_{OF} is the regularization parameter. The goal is to find a $\boldsymbol{w}$ which minimizes the above energy functional.

The multi-scale warping approach is integrated in the Euler-Lagrange equations of the above energy functional. It involves a downsampling operation, which allows linearisation of the grey value constancy assumption, thus leading to a linear system of equations. More specific details about the parameters and the optic flow method itself can be found in the paper by Brox et al. [4].

2.2 Evaluated SR Observational Models

In this section, we propose the super-resolution observational models that deviate from the image acquisition procedure mentioned in Equation 2. These Models 1-6 are specified in Table 1. The modifications that lead to the deviations from the imaging physics are a combination of dropping one of the blurring operators and permuting the operators. It should be mentioned that the warping and blurring matrices are space variant and hence none of the three operators (blur, warp and downsample) commute with each other. This also assures that none of the proposed models are equivalent to each other. The aim is to test which of the models gives the best SR reconstruction results.

For the evaluation of different observational models, we embed them in a variational framework. As mentioned in Section 1, we capitalized on the flexibility of the variational models for evaluating different data terms. Let us consider Model 1. The energy that has to minimized in order to obtain an SR image using the standard observational model (Model 1) is given by

$$E(\boldsymbol{f}_H) = \frac{1}{2}\sum_{k=1}^{N} |\boldsymbol{DBW}_k\boldsymbol{f}_H - \boldsymbol{f}_L^k|^2 + \frac{1}{2}\alpha|\boldsymbol{A}_1\boldsymbol{f}_H|^2 \tag{5}$$

where $\boldsymbol{A}_1$ is a discrete approximation of the gradient. A higher value of α would give rise to an SR image with a smoother gradient. The Euler-Lagrange equation of this energy functional is given by

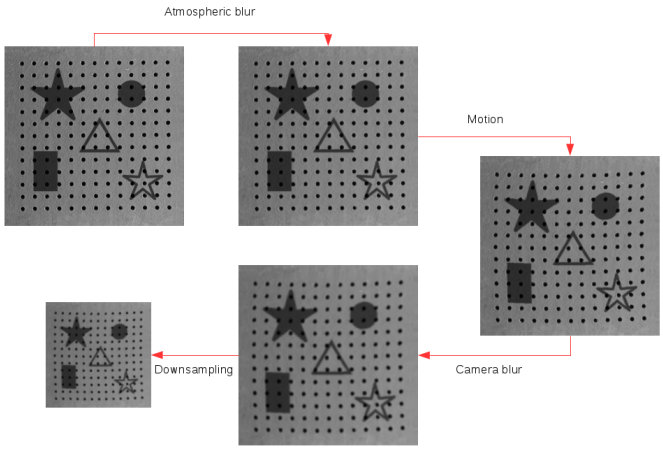


Fig. 1: Generation of Test Set 1 (TS1).

$$\sum_{k=1}^{N} \boldsymbol{W}_k^T \boldsymbol{B}^T \boldsymbol{D}^T (\boldsymbol{D}\boldsymbol{B}\boldsymbol{W}_k \boldsymbol{f}_H - \boldsymbol{f}_L^k) - \alpha |\boldsymbol{A}_2 \boldsymbol{f}_H| = 0 \tag{6}$$

where $\boldsymbol{A}_2$ is the discrete approximation of the Laplacian. We have used a Gaussian blurring kernel throughout this paper to model the blurring operator $\boldsymbol{B}$. Since $\boldsymbol{B}$ describes a space invariant Gaussian blur, the transposed matrix $\boldsymbol{B}^T$ is the same as $\boldsymbol{B}$. The matrix $\boldsymbol{D}^T$ is the upsampling operator while $\boldsymbol{D}$ is the downsampling operator. For mathematical consistency, $\boldsymbol{D}^T$ and $\boldsymbol{D}$ should be implemented with the same interpolation technique. We have used the area based interpolation technique for upsampling and downsampling processes. The matrix $\boldsymbol{W}_k$ represents the forward warping matrix while $\boldsymbol{W}_k^T$ denotes backward warping. We have implemented warping using bilinear interpolation.

We use a gradient descent scheme with parameters τ (the time step size) and ℓ_{max} (the number of iterations) to solve the Euler-Lagrange equations of all the proposed models in Table 1. The gradient descent scheme applied to the Euler-Lagrange equation of the standard model is given by

$$\boldsymbol{f}_H^{\ell+1} = \boldsymbol{f}_H^{\ell} + \tau \Big(\alpha \boldsymbol{A}_2 \boldsymbol{f}_H^{\ell} - \sum_{k=1}^{N} \boldsymbol{W}_k^T \boldsymbol{B}^T \boldsymbol{D}^T (\boldsymbol{D}\boldsymbol{B}\boldsymbol{W}_k \boldsymbol{f}_H^{\ell} - \boldsymbol{f}_L^k) \Big). \tag{7}$$

3 Experiments

3.1 Image Datasets

To evaluate the performance of the methods mentioned in Table 1, we have generated two LR test image sequences. More specifically, we have simulated the

(a) Original Image (b) Area upsampled reference frame (c) Model 1 (MSE: 103.26)

Fig. 2: SR reconstructed image of TS1 using F1 as optic flow.

image acquisition process of a HR scene on two ground truth HR images. The *shapes* image has a texture background taken from the texture image database of the Massachusetts Institute of Technology which is available at the following link: http://vismod.media.mit.edu/vismod/ imagery/VisionTexture/vistex.html. We refer to the 'shapes' image sequence as Test Set 1 (TS1). The second image sequence which is composed of self generated text is referred to as Test Set 2 (TS2). The ground truth of both image sequences is a 512×512 greyscale image. Figure 1 shows the generation of the image sequence TS1 from its ground truth. We have simulated a deformation type of motion with sub-pixel displacements using a combination of sine and cosine waves with randomly selected amplitude. It is well known that sub-pixel displacements are a requirement for super-resolution [17]. The ground truth image in TS2 undergoes a similar degradation process but without atmospheric blur. A zooming factor of $z = 3$ was used in both image datasets.

3.2 Parameters

Optic flow parameters: The model parameters are α_{OF} (smoothness/ regularization parameter for optic flow) and σ_{OF} (Gaussian pre-smoothing standard deviation). We have assumed a scenario where the ground truth optic flow is not available (which is generally the case) for selecting the optic flow parameters. Hence, we select sub-optimal optic flow model parameters by confirming visually that the images are properly registered. The numerical parameters are η (downsampling factor), η_1 (we use 10 inner fixed point iterations), η_2 (10 outer fixed point iterations) and ω (we select 1.95 for the SOR over-relaxation parameter). For evaluation purposes, we consider the same optic flow parameters for different data terms used for a particular image dataset.

SR parameters: We have optimized the two model parameters α (for regularization) and σ (standard deviation of the Gaussian kernel for blurring) with respect to the mean squared error (MSE) as the ground truth SR image is available. This helps in evaluating the performance of different SR data terms. The numerical parameters are ℓ_{max} and τ. A decay of the norm of the residue by a factor of 10^{-5} was used as the stopping criterion for iterations. We also utilize a Fast Explicit Diffusion (FED) scheme [22] to accelerate the explicit gradient

Table 2: Proposed SR observational models along with the MSE values of the obtained reconstructed SR image for TS2 using both ground truth (GT) and sub-optimal flow (SOF). The optimized parametric values are also shown.

#	Observational Model	MSE (GT)	MSE (SOF)	σ (GT)	σ (SOF)	α (GT)	α (SOF)
1	$\boldsymbol{DBW}_k\boldsymbol{f}_H = \boldsymbol{f}_L^k$	**10.85**	**173.18**	0.34	0.42	0.0002	0.003
2	$\boldsymbol{DW}_k\boldsymbol{B}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	**23.42**	**162.86**	0.35	0.45	0.0005	0.002
3	$\boldsymbol{BDW}_k\boldsymbol{f}_H = \boldsymbol{f}_L^k$	77.92	248.38	0.50	0.55	0.0009	0.005
4	$\boldsymbol{W}_k\boldsymbol{DB}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	250.35	294.00	0.31	0.33	0.001	0.0006
5	$\boldsymbol{BW}_k\boldsymbol{D}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	422.82	451.70	0.42	0.43	0.002	0.001
6	$\boldsymbol{W}_k\boldsymbol{BD}\boldsymbol{f}_H = \boldsymbol{f}_L^k$	423.52	451.91	0.42	0.43	0.002	0.002

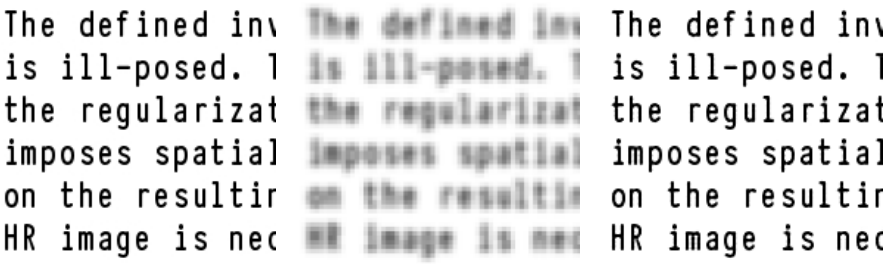
descent scheme. We observed that $\tau = 0.1$ was a stable time step size for all the proposed observational models through backtracking search.

It should be mentioned that modifications of the variational SR reconstruction method we used are definitely possible. For better results, we can use more complex optic flow models, anisotropic blurring kernels, robust data terms, discontinuity-preserving anisotropic smoothness terms and more sophisticated interpolation strategies. However, the aim of our work is to compare the performance of the data terms. To this very end, we keep things simple by assuming a noise-free scenario. For comparison purposes, we have integrated all our models on a common platform. It uses the same variational framework and default parameter settings for the optic flow algorithm.

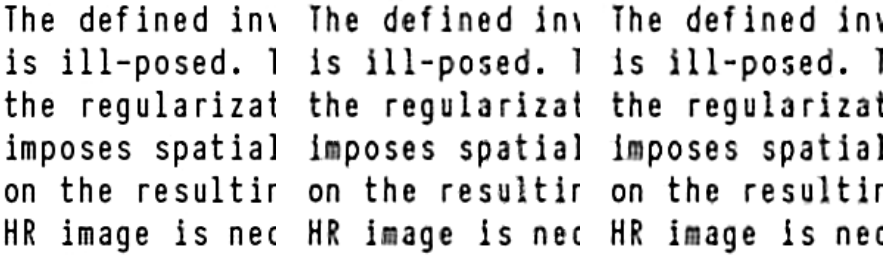
Both image sequences had 13 images and we have used the last image as the reference image in both cases. An area upsampled reference image was used as an initialisation for the solution of the gradient descent scheme.

3.3 Results and Discussion

As explained in Section 3.2, we have selected the optic flow parameters by confirming that the images are properly registered, assuming the real world scenario where the ground truth flow is not known. To emphasize the importance of a good optic flow for super-resolution, we have performed the experiments on TS1 using two different sets of optic flow parametric values. The first set of optic flow parameters with $\eta = 0.5, \sigma_{OF} = 0.3, \alpha_{OF} = 8.0$ is represented as F1. The optic flow parametric set $\eta = 0.5, \sigma_{OF} = 0.3, \alpha_{OF} = 30.0$ is denoted by F2. We have selected the parametric values of F2 such that the corresponding image registration was inferior when compared to the image registration with F1. It should be remembered that TS1 is generated by simulating both atmospheric blur and camera blur. From Table 1, we can conclude two things from the reported MSE values. Firstly, the best SR reconstructed image using F1 is better than the best SR reconstructed image using F2. This is not a surprising result since better optic flow leads to better image registration and hence a better reconstructed



(a) Original Image (b) Area upsampled reference image (c) Model 1 (GT, 10.85)

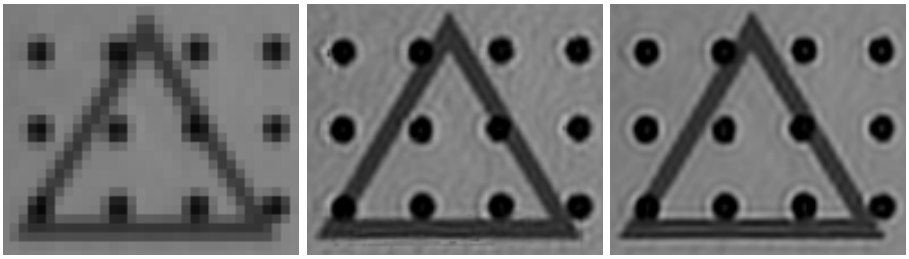


(d) Model 2 (GT, 23.42) (e) Model 1 (SOF, 173.18) (f) Model 2 (SOF, 162.86)

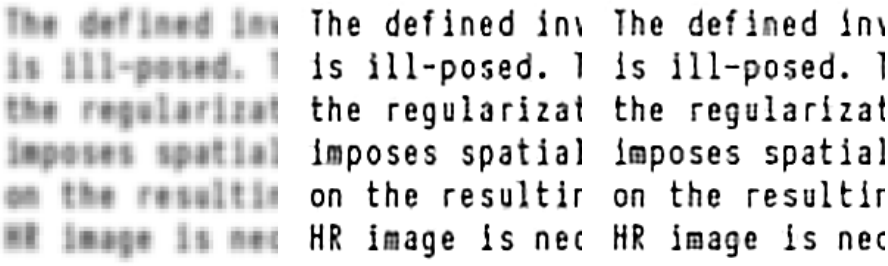
Fig. 3: SR reconstructed images of TS2 along with MSE and flow.

SR image. Secondly, the ranking of the data terms with respect to the MSE value has changed with change in optic flow. This is more surprising than the first conclusion. With a better optic flow, the standard SR observational model (Model 1) gives the best results by a fair margin. If the optic flow is bad (F2), this is no longer the case. Thus, this experiment emphasizes the importance of optic flow in both quality of the reconstructed image and the ranking of the data terms. Figure 2 shows the reconstructed image of TS1 using F1 and Model 1.

From the above discussion, we could think that, to conclude correctly about the ranking of the proposed SR observational models in Table 1, the ground truth optic flow has to be used. Hence for TS2, we have used both ground truth (GT) optic flow and sub-optimal optic flow (SOF) with $\eta = 0.9, \alpha_{OF} = 15.0, \sigma_{OF} = 0.3$. One needs to remember that TS2 has been generated without simulating atmospheric blur. This is the general case scenario where camera blur is dominant over atmospheric blur. Table 2 shows the MSE values of the reconstructed SR image of TS2 while Figure 3 shows the reconstructed images. We can conclude that the standard observational model gives the best results for the ground truth optic flow. This is similar to the observation in the previous experiment where Model 1 gave the best results for better optic flow F1. On using SOF as the optic flow parametric values, Model 2 gave the best results. It can be observed that the error in the best reconstructed image using GT is much smaller than the error in the best reconstructed image using SOF. This reinforces the critical nature



(a) Area upsampled reference image (b) Model 1 (F1, 103.26) (c) Model 2.1 (F1, 97.05)



(d) Area upsampled reference image (e) Model 2 (SOF, 162.86) (f) Model 2.1 (SOF, 172.03)

Fig. 4: SR reconstructed images of TS1 and TS2 using Model 2.1.

of motion estimation in SR reconstruction. The main reason behind Models 1 and 2 performing better is that they are the closest to the image acquisition model in Equation 2. In other words, the only manipulation they undergo while they are derived from the image acquisition model is dropping one of the blur operators but they do not undergo any swapping of operators like in Models 3-6. The lesson to be learnt from this experiment is that the observational model needs to stay as close as possible to the imaging physics. Now, is this a trivial result? We will be answering this question shortly.

Zhang et al. [25] and Rockefort et al. [18] have discussed the application of both observational Models 1 and 2 but only when affine motion is assumed. Wang et al. [21] provide a short but insightful work. They also discuss Models 1 and 2 in the case where camera blur is dominant over atmospheric blur. They do not have constraints on the type of motion. With the help of simple counter arguments based on the theory of operator commutability, they prove that, in the case where motion has to be estimated from the LR images (similar to the case where sub-optimal flow has been used for TS2), SR reconstruction using Model 1 introduces a systematic error. Model 2 does not introduce this error and hence it should be preferred over Model 1. The other case is the one where the motion between the HR scenes is known (similar to the case where we have used ground truth flow). In such a case Model 1 gives better results than Model 2 as there is no systematic error. Wang et al. also support their arguments with experimental validations. One can see that by observing the MSE values

and the reconstructed images of TS2, the experimental validations of [21] are confirmed. However, by using a multi-scale warping approach, we have employed a more advanced variational motion estimation method than the one [3] by Wang et al. [21]. This still doesn't guarantee that the standard observational model performs better. Also, they discard the smoothness term in evaluating Models 1 and 2. Thus, with our experiments, we can additionally conclude that inclusion of a regularization term does not change the ranking of the models with respect to SR reconstruction. One could expect to reduce this performance gap between the two Models 1-2 by using even better optic flow.

Thus, going back to the unanswered question above, it is not confirmed that if we deviate from the image acquisition process but still follow a mathematically plausible observational model, we are bound to get worse results. The arguments in the above paragraph support this statement. In such a case it is really necessary that one verifies all mathematically plausible models. Thus, our contribution of evaluating all the six models is not a trivial experiment. However, it turned out that Models 3-6 are outperformed by Models 1-2.

Eventhough Model 2 does not outperform Model 1 by a large margin (while using sub-optimal flow), when we retrospect the reported works on super-resolution after [21] was reported, it is surprising that in most of the works the standard observational model has been used [1, 2, 10, 11, 14, 15, 24]. Model 2 has not been considered as a possible observational model.

3.4 More Efficient Model

Now we propose another mathematically plausible model derived from Model 2 and discuss what could be the advantages of using it. The following is the representation of what we denote as Model 2.1: $\boldsymbol{B}\boldsymbol{f}_H = \boldsymbol{W}_k^T \boldsymbol{D}^T \boldsymbol{f}_L^k$. It is clear that it is dervied from Model 2 as the ordering of operators is the same. However, it is definitely different from Model 2 itself as warping and downsampling are interpolation operations. Interpolation in general is not an invertible operation. This model also deviates from the imaging physics but is still a mathematically plausible model. The gradient descent of the Euler-Lagrange equation corresponding to the energy using Model 2.1 in the data term, is given by:

$$\boldsymbol{f}_H^{\ell+1} = \boldsymbol{f}_H^{\ell} + \tau(\alpha \boldsymbol{A}_2 \boldsymbol{f}_H^{\ell} - (N \boldsymbol{B}^T \boldsymbol{B} \boldsymbol{f}_H^{\ell} - \boldsymbol{C})) \tag{8}$$

where $\boldsymbol{C} = \sum_{k=1}^{N} \boldsymbol{B}^T \boldsymbol{W}_k^T \boldsymbol{D}^T \boldsymbol{f}_L^k$ can be precomputed. Such a precomputation is not possible with Model 2.

Figure 4 shows the reconstructed images using Model 2.1. The parameters $\sigma = 0.64, \alpha = 0.006$ were selected for TS1 and $\sigma = 0.59, \alpha = 0.002$ were selected for TS2 after optimizing for the MSE with respect to the ground truth SR image. We can conclude from the Figure 4 that the reconstructed HR images obtained using Model 2.1 are not far off from Models 1 and 2 in terms of image reconstruction quality. However, it was observed that because of being able to pre-compute $\boldsymbol{C}$, the gradient descent of Model 2.1 is twenty-five times faster for TS1 and eighteen times faster for TS2 when compared to the gradient descent

schemes of Models 1 and 2 respectively. This can be a decisive advantage in time critical applications.

4 Conclusion and Outlook

Super-resolution requires to model three physical penomena: blur, warp, and downsample. In our paper, we have performed the first systematic evaluation of the influence of the order of these three operators on the result of a variational super-resolution model. This has led to the surprising result that it is not always the physically most plausible and most widely used model which performs best in a practical setting when motion estimation may suffer from errors. Thus, it is worthwhile to consider also alternative models. Moreover, we saw that closely related models can lead to algorithms with strongly differing efficiency: By reformulating the blur-warp-downsample model we managed to come up with a novel model that was 18–25 times more efficient. These insights emphasize the fundamental importance of careful model design.

A possible future work would be to confirm the evaluation performed in this paper also for the case of color images. We are also going to apply our super-resolution algorithms to real world images from biophysics that suffer from severe noise and the absence of ground truth data.

References

1. Babacan, S., Molina, R., Katsaggelos, A.: Variational Bayesian super resolution. IEEE Transactions on Image Processing 20(4), 984–999 (Apr 2011)
2. Belekos, S., Galatsanos, N., Katsaggelos, A.: Maximum a posteriori video super-resolution using a new multichannel image prior. IEEE Transactions on Image Processing 19(6), 1451–1464 (Jun 2010)
3. Black, M., Anandan, P.: A framework for robust estimation of optical flow. In: Proc. Fourth International Conference on Computer Vision. pp. 231–236. IEEE Computer Society Press, Berlin, Germany (May 1993)
4. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In: Pajdla, T., Matas, J. (eds.) Computer Vision – ECCV 2004, Part IV, Lecture Notes in Computer Science, vol. 3024, pp. 25–36. Springer, Berlin (2004)
5. Chaudhuri, S.: Super-resolution imaging. Springer, New York (2001)
6. Elad, M., Feuer, A.: Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. IEEE Transactions on Image Processing 6(12), 1646–1658 (Dec 1997)
7. Elad, M., Hel-Or, Y.: A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. IEEE Transactions on Image Processing 10(8), 1187–1193 (Aug 2001)
8. Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multi-frame super resolution. IEEE Transactions on Image Processing 13(10), 1327–1364 (Oct 2004)

9. Glasner, D., Shai, B., Irani, M.: Super-resolution from a single image. In: Proc. Twelfth International Conference on Computer Vision. pp. 349–356. IEEE, Kyoto, Japan (Sep 2009)
10. He, H., Kondi, L.: An image super-resolution algorithm for different error levels per frame. IEEE Transactions on Image Processing 15(3), 592–603 (Mar 2006)
11. Li, X., Hu, Y., Gao, X., Tao, D., Ning, B.: A multi-frame image super-resolution method. Signal Processing 90(2), 405–414 (Feb 2010)
12. Marquina, A., Stanley, J.: Image super-resolution by TV-regularization and Bregman iteration. Springer Journal of Scientific Computing 37(3), 367–382 (Jul 2008)
13. Milanfar, P.: Super-resolution Imaging. CRC Press, Boca Raton (Sep 2010)
14. Mitzel, D., Pock, T., Schoenemann, T., Cremers, D.: Video super resolution using duality based TV-l1 optical flow. In: Denzler, J., Notni, G., Süße, H. (eds.) Pattern Recognition, Lecture Notes in Computer Science, vol. 5748, pp. 432–441. Springer, Berlin (Sep 2009)
15. Ng, M., Shen, H., Lam, E., Zhang, L.: A total variation regularization based super-resolution reconstruction algorithm for digital video. EURASIP Journal on Advances in Signal Processing (1), 1–6 (Dec 2007)
16. Nguyen, N., Milanfar, P., Golub, G.: Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement. IEEE Transactions on Image Processing 10(9), 1299–1308 (Sep 2001)
17. Park, S., Park, M., Kang, M.: Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine 20(3), 21–36 (May 2003)
18. Rochefort, G., Champagnat, F., Besnerais, G.L., Giovannelli, J.: An improved observation model for super-resolution under affine motion. IEEE Transactions on Image Processing 15(11), 3325–3337 (Nov 2006)
19. Tian, J., Ma, K.: A survey on super-resolution imaging. Signal, Image and Video Processing 5(3), 329–342 (Sep 2011)
20. Tsai, R., Huang, T.: Multiframe image restoration and registration. Advances in Computer Vision and Image Processing 1(2), 317–339 (Nov 1984)
21. Wang, Z., Qi, F.: On ambiguities in super-resolution modeling. IEEE Signal Processing Letters 11(8), 678–681 (Aug 2004)
22. Weickert, J., Grewenig, S., Schroers, C., Bruhn, A.: Cyclic schemes for PDE-based image analysis. International Journal of Computer Vision 118(3), 275–299 (Jul 2016)
23. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing 19(11), 2681–2873 (May 2010)
24. Zhang, L., Zhang, H., Shen, H., Li, P.: A super-resolution reconstruction algorithm for surveillance images. Signal Processing 90(3), 848–859 (Mar 2010)
25. Zhang, X., Jiang, J., Peng, S.: Commutability of blur and affine warping in super-resolution with application to joint estimation of triple-coupled variables. IEEE Transactions on Image Processing 21(4), 1796–1808 (Apr 2012)